

## INICIATIVA ACADÉMICA DE VISIÓN POR COMPUTADORA

### 1 Identificación

<b>Curso:</b>	CC3102 – Visión por Computadora	<b>Créditos:</b>	4
<b>Ciclo:</b>	Primero	<b>Requisitos:</b>	Algoritmos y Estructuras de Datos Deep Learning Álgebra Lineal
<b>Año:</b>	2024		
<b>Profesor:</b>	Alan Reyes-Figueroa	<b>Horario:</b>	Martes 18:10-19:45 y jueves 17:20-18:55
<b>Email:</b>	agreyes	<b>Sala:</b>	H-301 y CIT-312.

#### Sitio Web del Curso:

- <https://pfafer.github.io/cv2024>

#### Office Hours:

- Por solicitud del estudiante, o pueden enviar sus dudas por correo electrónico.

### 2 Descripción

Este es un curso introductorio a la visión computacional (CV). Inicialmente, haremos una revisión de las técnicas estándar usadas en el procesamiento de imágenes digitales, el diseño de filtros básicos y aplicaciones de transformaciones, las cuales sirven como base para el desarrollo de aplicaciones inteligentes asociadas a imágenes.

Aprenderemos y estudiaremos los algoritmos más comunes para la detección, extracción y comparación de características. Abordaremos también el estudio geométrico de imágenes de una vista (one-point-view) y de dos vistas (two-point-view), y sus transformaciones asociadas. Estos algoritmos se utilizan para alinear y unir imágenes para crear una única imagen de una escena más grande, y para recrear escenas 3D a partir de imágenes planas.

Abordaremos también grandes temas como la clasificación y segmentación de imágenes, y estudiaremos los métodos actuales de aprendizaje automático para este objetivo, principalmente redes neuronales convolucionales. Abordaremos también temas como la detección de objetos en una imagen, así como estimación de movimiento y seguimiento de objetos con aprendizaje automático. Al final del curso y si el tiempo lo permite haremos un breve recorrido por metodologías y herramientas actuales de IA generativa, para producir imágenes a partir de descripciones.

El curso requiere madurez por parte del estudiante, pues se integran contenidos de muchos cursos de computación, matemática y estadística. Entre los prerrequisitos se encuentra tener un buen dominio de las técnicas vistas en los cursos de matemática discreta, grafos, álgebra lineal, programación y algoritmos, cálculo diferencial e integral, y estadística.

### 3 Competencias a Desarrollar

#### Competencias genéricas

1. Piensa de forma crítica y analítica.
2. Resuelve problemas de forma efectiva.
3. Desarrolla habilidades de investigación y habilidades de comunicación científica a través de seminarios y presentaciones ante sus colegas.

#### Competencias específicas

- 1.1 Identifica y conoce los aspectos fundamentales del procesamiento de imágenes y la visión computacional, para tener una visión global de los orígenes y motivaciones de ésta área.
- 1.2 Distingue el concepto de cámara, modelo de cámara, transformaciones proyectivas, filtro digital, convolución discreta, morfología, puntos característicos, alineación, entre otros.
- 1.3 Comprende y conoce la terminología común en las áreas de procesamiento de imágenes, visión por computadora, y redes neuronales convolucionales.
- 2.1 Evalúa correctamente los métodos más adecuados para abordar un problema relacionado con imágenes digitales, vídeos, así como sus usos prácticos, tomando en consideración limitantes de disponibilidad de datos.
- 2.2 Construye algoritmos específicos de visión por computadora para resolver eficientemente problemas computacionales asociados a imágenes o vídeo.
- 2.3 Utiliza un enfoque global para resolver problemas. Utiliza herramientas auxiliares en su solución, como álgebra lineal, estadística y ciencia de datos, lógica y algoritmos.
- 3.1 Desarrolla todas las etapas de un proyecto aplicado donde se realiza una implementación de algoritmos de visión computacional.
- 3.2 Escribe un reporte técnico sobre la solución de un problema aplicado. Concreta un análisis riguroso y conclusiones importantes.
- 3.3 Comunica de manera efectiva, en forma escrita, oral y visual, los resultados de su investigación.

### 4 Metodología Enseñanza Aprendizaje

El curso se desarrollará durante diecinueve semanas, con cuatro períodos semanales de cuarenta y cinco minutos para desenvolvimiento de la teoría, la resolución de ejemplos y problemas, comunicación didáctica y discusión. Se promoverá el trabajo colaborativo de los estudiantes por medio de listas de ejercicios.

El resto del curso promoverá la revisión bibliográfica y el auto aprendizaje a través de la solución de los ejercicios del texto, y problemas adicionales, y el desarrollo de una monografía. Se espera que el estudiante desarrolle su trabajo en grupo o individualmente, y que participe activamente y en forma colaborativa durante todo el curso.

## 5 Contenido

1. Procesamiento de imágenes: (6 semanas) Imágenes digitales. Principales formatos y tipos de imágenes. Formatos cromáticos y conversión entre ellos. Filtros y convolución. Diseño y efectos de filtros. Gradientes y líneas. Detección de bordes. Aplicaciones en restauración de imágenes.

Operadores morfológicos: opening, closing, top-hat. Transformaciones morfológicas. Esqueletización. Detección de componentes. Histogramas y ecualización. Algoritmos para segmentación binaria de una imagen.

Filtros de Gabor, bancos de filtros. Transformada de Fourier, transformada de Hough. Submuestreo y efectos, aliasing y artefactos. Eliminación de artefactos. Pirámides. Filtros de Haar. Aplicaciones a detección de objetos.

2. Geometría de 1 y 2 vistas. (3 semanas) Transformaciones proyectivas y homografías. Modelos de cámaras. Matriz fundamental. Calibración de cámaras. Geometría epipolar, visión estéreo y profundidad.
3. Física y óptica de las imágenes. (1 semana) Formación de imágenes físicas y digitales. Cámaras y lentes. Color, luz y aberraciones ópticas. Espectro visible. Sensores biológicos y sensores digitales.
4. Alineación de imágenes. (2 semanas) Puntos característicos, Descriptores. Detección de esquinas. Algoritmos principales: SIFT, SURF. RANSAC. Aplicaciones para matching y stitching de imágenes. Construcción de panoramas.
5. Aprendizaje automático aplicado a imágenes. (6 semanas) Redes convolucionales (CNN o ConvNets). Tips para calibración y entrenamiento. Arquitecturas para clasificación de imágenes: AlexNet, ResNet, GoogleNet. Bag of Words y descriptores. Algoritmos de clasificación no-supervisada:  $K$ -means *et al.*

Arquitecturas para segmentación de imágenes: U-Net, Mask R-CNN. Otros algoritmos para segmentación. Detección de objetos. Arquitecturas para detección de objetos: Regional-CNN, ResNet-50, YOLO, YOLO2.

Estimación de movimiento y Flujo óptico: algoritmo de Horn-Schunk. Tracking y seguimiento de objetos: algoritmo de Lucas-Kanade.

GANS, IA generativa y métodos de difusión, aplicaciones y herramientas actuales para generación de imágenes (StableDiffusion, MidJourney, DaLL-e).

## 6 Bibliografía

### Textos:

- R. Szeliski (2022). *Computer Vision: Algorithms and Applications*.

### Referencias adicionales:

- R. Hartley y A. Zisserman (2009). *Multiple View Geometry in Computer Vision*.
- D. Forsyth y J. Ponce (2000). *Computer Vision: A Modern Approach*.
- R. Gonzalez y R. Woods (1999). *Digital Image Processing*.

## 7 Actividades de evaluación

Actividad	Cantidad aproximada	Porcentaje
Listas y Laboratorios	6	20%
Proyectos	4	80%

## 8 Cronograma

Semana	Tópico	Fecha	Actividades
1	Conceptos básicos. (Formación de imágenes, luz, lentes, color, receptores)	08-12 enero	
2		15-19 enero	
3	Procesamiento de imágenes. (Filtros, bordes, morfología, threshold, procesamiento)	22-26 enero	
4		29 enero-02 febrero	
5		05-09 febrero	
6	Transformadas. (Pirámides, Fourier, transformada Hough, filtros Haar)	12-16 febrero	
7		19-23 febrero	
8		26 febrero-01 marzo	
9	Descriptores. (SIFT, SURF, Ransac, Matching)	04-08 marzo	
10		11-15 marzo	
11	Geometría de 1 y 2 vistas. (Transformaciones 2D, homografías, modelos de cámaras)	18-22 marzo	
	<i>Semana Santa</i>	25-29 marzo	
12		01-05 abril	
13		08-12 abril	
14		15-19 abril	
15	Visión en la era del Deep Learning. (Redes neuronales, Redes convolucionales)	22-26 abril	
16	(clasificación imágenes, segmentación)	29 abril-03 mayo	
17	(detección objetos)	06-10 mayo	
18	(tracking, restoration, style transfer, ...)	13-17 mayo	
19	(aplicaciones con IA generativa)	20-24 mayo	
20	Presentación de proyectos.	27-31 mayo	